Security Based Distributed Data Storage By Using Cloud Computing

Dr.T.Chalama Reddy [*1] T.Sampath Reddy[*2] Peddadoddi Raju[*3]
[*1]Professor, Department of Computer Science and Engineering
[*2]AssistantProfessor, Department of Computer Science and Engineering
[*3]M-Tech Student, Department of Computer Science and Engineering
[*1,2,3]Nalla Narshima Reddy Group of Institution, Narapally.

***Abstract*— Distributed computing has increased extraordinary consideration from both industry and the scholarly community since 2007.With the objective of giving clients increasingly adaptable administrations in a straightforward way, all administrations are allotted in a "cloud" that really is an accumulation of gadgets and assets associated through the Internet. One of the center administrations given by distributed computing is information stockpiling. This postures new difficulties in making secure and solid information stockpiling and access offices over remote administration suppliers in the cloud. The security of information stockpiling is one of the fundamental errands to be tended to before the outline for distributed computing is acknowledged we have exhibited the cutting edge look into advancement and results of secure circulated information stockpiling in distributed computing. Distributed computing has gained impressive consideration from both industry and the scholarly world lately. Among all the major building squares of distributed computing, information stockpiling plays a very vital job. At present, there are a few difficulties in executing conveyed capacity in distributed computing situations. These difficulties should be tended to before clients can appreciate the full favorable circumstances of distributed computing. In expansion, security is dependably a noteworthy issue in any figuring framework. Subsequently, we overviewed a number of subjects related to the testing issues of verifying appropriated information stockpiling, including database redistributing and question trustworthiness affirmation, information uprightness in conniving stockpiling, Web-application-based security, what's more, interactive media information security. It is foreseen that the advancements created in the previously mentioned research will add to making ready for verifying dispersed information stockpiling situations inside distributed computing.**

Keywords*— *Distributed computing, Cloud Storage, Cloud services, Security.

I. INTRODUCTION

Distributed computing has increased extraordinary consideration from both industry and the scholarly community since 2007.With the objective of giving clients increasingly adaptable administrations in a straightforward way, all administrations are allotted in a "cloud" that really is an accumulation of gadgets and assets associated through the Internet. One of the center administrations given by distributed computing is information stockpiling. This postures new difficulties in making secure and solid information stockpiling and access offices over remote administration suppliers in the cloud. The security of information stockpiling is one of the fundamental errands to be tended to before the outline for distributed computing is acknowledged.

In the previous decades, information stockpiling has been perceived as one of the primary worries of data innovation. The advantages of system based applications have prompted the change from server-appended capacity to appropriated capacity. In view of the way that information security is the establishment of data security, an extraordinary amount of endeavors has been made in the region of disseminated stockpiling security [13]. Be that as it may, this exploration in distributed computing security is still in its early stages [4].

One thought is that the novel issues related with distributed computing security have not been perceived. A few scientists imagine that distributed computing security won't be vastly different from existing security rehearses and that the security angles can be all around overseen utilizing existing procedures, for example, advanced marks, encryption, firewalls, as well as the segregation of virtual conditions, etc [4]. For instance, SSL (Secure Sockets Layer) is a convention that gives solid secure correspondences on the Internet for things, for example, Web perusing, email, texting, and other information exchanges.

Present technologies for data security in cloud computing from four different perspectives:

1 Database Outsourcing and Query Integrity Assurance
2 Data Integrity in Untrustworthy Storage
3 Web-Application-Based Security
4 Multimedia Data Security Storage

II. CLOUD STORAGE: FROM LANS TO WANS

It will be a revolutionary change in computing services. Users will be allowed to purchase CPU cycles, memory utilities, and information storage services conveniently just like how we pay our monthly water and electricity bills. However, this image will not become realistic until some challenges have been addressed. In this section, we will briefly introduce the major difference brought by distributed data storage in cloud computing environment. Then, vulnerabilities in today's cloud computing platforms are analyzed and illustrated. Most designs of distributed storage take the form of either storage area networks (SANs) or network-attached storage (NAS) on the LAN level, such as the networks of an enterprise, a campus, or an organization. SANs are constructed on top of block-addressed storage units connected through dedicated high-speed networks. In

contrast, NAS is implemented by attaching specialized file servers to a TCP/IP network and providing a file-based interface to client machine [6]. For SANs and NAS, the distributed storage nodes are managed by the same authority. The system administrator has control over each node, and essentially the security level of data is under control.

The reliability of such systems is often achieved by redundancy, and the storage security is highly dependent on the security of the system against the attacks and intrusion from outsiders. The confidentiality and integrity of data are mostly achieved using robust cryptographic schemes. However, such a security system would not be robust enough to secure

The data in distributed storage applications at the level of wide area networks, specifically in the cloud computing environment. The recent progress of network technology enables global-scale collaboration over heterogeneous networks under different authorities. For instance, in a peer-to-peer (P2P) file sharing environment, or the distributed storage in a cloud computing environment, the specific data storage strategy is transparent to the user Further more, there is no approach to guarantee that the data host nodes are under robust security protection. In addition, the activity of the medium owner is not controllable to the data owner. Theoretically speaking, an attacker can do whatever she wants to the data stored in a storage node once the node is compromised. Therefore, the confidentiality and the integrity of the data would be violated when an adversary controls a node or the node administrator becomes malicious.

A. Existing Commercial Cloud Services

Information stockpiling administrations on the stage of distributed computing are essentially given by applications/programming dependent on the Internet. Despite the fact that the meaning of distributed computing isn't clear yet, a few pioneer business usage have been developed and opened to people in general, for example, Amazon's Computer Cloud AWS (Amazon Web administration) [7], the Microsoft Azure Service Platform [8], and the Google App Engine (GAE) [9]. In ordinary system based applications, client verification, information privacy, furthermore, information honesty can be unraveled through IPSec intermediary utilizing encryption furthermore, advanced mark. The key trading issues can be unraveled by SSL intermediary. These techniques have been connected to the present distributed computing to verify the information on the cloud and furthermore secure the correspondence of information to and from the cloud. The specialist co-ops guarantee that their administrations are secure. This segment portrays three secure strategies utilized in three business cloud administrations also, talks about their vulnerabilities

Amazon's Web Service.

Amazon provides Infrastructure as a Service (IaaS) with different terms, such as Elastic Compute Cloud (EC2), Simple(DB), Simple Storage Service (S3), and so on. They are supposed to ensure the confidentiality, integrity, and availability of the customers' applications and data. Figure presents one of the data processing methods adopted in Amazon's AWS ,which is used to transfer large amounts of data between the AWS cloud and portable storage devices When the user wants to upload the data, he/she stores some parameters such as AccessKeyID, DeviceID, Destination, and so on, into an import metadata file called the manifest file and then signs the manifest file and e-mails the signed manifest file to Amazon. Another metadata file named the signature file is used by AWS to describe the cipher algorithm that is adopted to encrypt the job ID and the bytes in the manifest file. The signature file can uniquely identify and authenticate the user request. The signature file is attached with the storage device, which is shipped to Amazon for efficiency. On receiving the storage device and the signature file, the service provider will validate the signature in the device with the manifest file sent through the email. Then, Amazon will e-mail management information back to the user including the number of bytes saved, the MD5 of the bytes, the status of the load, and the location on the Amazon S3 of the AWS Import Export Log. This log contains details about the data files that have been uploaded, including the key names, number of bytes, and MD5 checksum values

Microsoft Windows Azure

The Windows Azure Platform (Azure) is an Internet-scale cloud services platform hosted in Microsoft data centers, which provides an operating system and a set of developer services that can be used individually or together [8]. The platform also provides scalable storage service .There are three basic data items: blobs (up to 50 GB), tables, and queues(8k). In the Azure Storage, based on the blob, table, and queue structures, Microsoft promises to achieve confidentiality of the users' data. The procedure shown in Figure provides security for data accessing to ensure that the data will not be lost.

To use Windows Azure Storage service, a user needs to create a storage account, which can be obtained from the Windows Azure portal web interface. After creating an account, the user will receive a 256-bit secret key. Each time when the user wants to send the data to or fetch the data from the cloud, the
user has to use his secret key to create a HMAC SHA256 signature for each individual request for identification. Then the user uses his signature to authenticate request at server. The signature is passed with each request to authenticate the user requests by verifying the HMAC signature

Google App Engine (GAE)

The Google App Engine (GAE) [9] provides a powerful distributed data storage service that features a query engine and transactions. An independent third-party auditor, who claims that GAE can be secure under the SAS70 auditing industry standard, issued Google Apps an unqualified SAS70 Type II certification. However, from its on-line storage technical document of lower API [9], there are only some functions such as GET and PUT. There is no

content addressing the issues of securing storage services. The security of data storage is assumed guaranteed using techniques such as by SSL link, based on our knowledge of security method adopted by other service is one of the secure services, called Google Secure Data Connector (SDC), based on GAE [9]. The SDC constructs an encrypted connection between the data source and Google Apps. As long as the data source is in the Google Apps domain to the Google tunnel protocol servers, when the user wants to get the data, he/she will first send an authorized data requests to Google Apps, which forwards the request to the tunnel server. The tunnel servers validate the request identity. If the identity is valid, the tunnel protocol allows the SDC to set up a connection, authenticate, and encrypt the data that flows across the Internet. At the same time, the SDC uses resource rules to validate whether a user is authorized to access a specified resource.When the request is valid, the SDC performs a network request. The server validates the signed request, checks the credentials, and returns the data if the
user is authorized.

B. Vulnerabilities in Current Cloud Services

Previous subsections describe three different commercial cloud computing secure data storage schemes. Storage services that accept a large amount of data (.1 TB) normally adopt strategies that help make the shipment more convenient, just as the Amazon AWS does. In contrast, services that only
accept a smaller data amount (#50 GB) allow the data to be uploaded or downloaded via the Internet, just as the Azure Storage Service does. To provide data integrity, the Azure Storage Service stores the uploaded data MD5 checksum in the database and returns it to the user when the user wants to
retrieve the data. Amazon AWS computes the data MD5 checksum and e-mails it to the user for integrity checking. The SDC is based on GAE's attempt to strengthen Internet authentication using a signed request. If these services are
grouped together, the following scheme can be derived

1.**Confidentiality**. Eve is considered as an untrustworthy third party, Alice and Bob do not want reveal the data to Eve.

2.**Integrity**. As the administrator of the storage service,Eve has the capability to play with the data in hand. How can Bob be confident that the data he fetched from Eve are the same as what was sent by Alice?Are there any measures to guarantee that the data have not been tampered by Eve.

3. **Repudiation**. If Bob finds that the data have been tampered with, is there any evidence for him to demonstrate that it is Eve who should be responsible for the fault? Similarly, Eve also needs certain evidence to prove her innocence.

As of late, a potential client made an inquiry on a cloud mailing-gathering as to honesty and administration unwavering quality. The answer from the engineer was "We won't lose your information we have a strong reinforcement and recuperation procedure be that as it may, we're not in charge of you losing your own information clearly, it is not enticing to the potential client to be certain with the administration.

The disavowal issue opens an entryway for conceivably blackmailers when the client is malevolent. How about we expect that Alice needs to extort Eve. Eve is a cloud capacity specialist organization who asserts that information honesty is one of their key highlights. For that reason, Alice put away a few information in the cloud, and later she downloaded the information. At that point, she detailed that her information were wrong and that it is the blame of the capacity supplier. Alice claims pay for her socalled misfortune. By what method can the specialist organization show her blamelessness
Upload-to-Download Integrity. Since the integrity in uploading and downloading phase are handled separately, how can the user or provider know the data retrieved from the cloud is the same data that the user uploaded previously?
Repudiation Between Users and Service Providers
When data errors happen without transmission errors in the uploading and downloading sessions, how can the user and service provider prove their innocence

C. Bridge the Missing Link

This section presents several simple ideas to bridge the missing link based on digital signatures and authentication coding schemes. According to whether there is a third authority certified (TAC) by the user and provider and whether the user and provider are using the secret key sharing technique (SKS), there are four solutions to bridge the missing link of data integrity between the uploading and downloading procedures. Actually, other digital signature technologies can be adopted to fix this vulnerability with different approaches
Neither TAC nor SKS.

•Uploading Session
•Downloading Session

III. TECHNOLOGIES FOR DATA SECURITY IN CLOUD COMPUTING

This section presents several technologies for data security and privacy in cloud computing. Focusing on the unique issues of the cloud data storage platform, this section does not repeat the normal approaches that provide confidentiality ,integrity, and availability in distributed data storage applications. Instead, we select to illustrate the unique requirements for cloud computing data security from a few different perspectives.

• Database Outsourcing and Query Integrity Assurance
• Data Integrity in Untrustworthy Storage.
• Web-Application-Based Security
• Multimedia Data Security

A. Database Outsourcing and Query Integrity Assurance

As of late, database re-appropriating has turned into a critical part of distributed computing. Because of the quick progressions in system innovation, the cost of transmitting a terabyte of information over long separations has diminished altogether in the previous decade. What's more, the all out expense of information the executives is five to multiple times higher than the underlying procurement costs. Thus, there is a developing enthusiasm for re-appropriating database the board undertakings to outsiders that can give these undertakings to a much lower cost because of the economy of scale. This new re-appropriating model has the advantages of decreasing the expenses for running Database Management Systems (DBMS) freely and empowering endeavors to focus on their fundamental organizations [12]. Figure 8.7 exhibits the general engineering of a database re-appropriating condition with customers. The database proprietor re-appropriates its information the executives errands, and customers send questions to the untrusted specialist organization. Give T a chance to mean the information to be re-appropriated. The information T are is preprocessed, scrambled, and put away at the specialist organization. For assessing inquiries, a client revises a lot of questions Q against T to questions against the scrambled database

• Data Privacy Protection
• Query Integrity Assurance

B. Data Integrity in Untrustworthy Storage

While the straightforward cloud gives adaptable utility of system based assets, the dread of loss of control on their information is one of the real concerns that keep end clients from relocating to distributed storage administrations. As a matter of fact it is a potential hazard that the capacity foundation suppliers wind up self-intrigued, deceitful, or even noxious. There are diverse inspirations whereby a capacity specialist organization could wind up conniving—for example, to cover the result of an oversight in activity, or deny the defenselessness in the framework after the information have been stolen by an enemy. This segment presents two advances to empower information proprietors to confirm the information trustworthiness while the documents are put away in the remote conniving stockpiling administrations.

C. Web-Application-Based Security

In distributed computing conditions, assets are given as an administration over the Web in a dynamic, virtualized, and adaptable way [29, 30]. Through cloud processing administrations, clients get to business applications on-line from a Web program, while the product and information are put away on the servers. In this manner, in the period of distributed computing, Web security assumes a more imperative job than any other time in recent memory. The Web website server is the main entryway that watches the tremendous cloud assets. Since the cloud may work ceaselessly to process a great many dollars of day by day on-line exchanges, the effect of any Web security helplessness will be intensified at the dimension of the entire cloud.

• Authentication
• Authorization
• Client-Side Attacks
• Command Execution
• Information Disclosure
• Logical Attack

D. Multimedia Data Security Storage

With the quick advancements of sight and sound advances, to an ever increasing extent sight and sound substance are being over numerous sorts of gadgets, databases, and systems. Mixed media Data Security assumes a critical job in the information stockpiling to ensure sight and sound information. As of late, how stockpiling interactive media substance are conveyed by both diverse suppliers and clients has pulled in much considerations and numerous applications. put away and conveyed This area quickly experiences the most basic themes around there

• Protection from Unauthorized Replication
• Protection from Unauthorized Replacement
• Protection from Unauthorized Pre-fetching

IV. OPEN QUESTIONS AND CHALLENGES

Practically all the present business cloud specialist co-ops guarantee that their stages are secure and strong. On one hand, they receive vigorous figure calculations for secrecy of put away information; then again, they rely upon organize correspondence security conventions, for example, SSL, IPSec, or others to secure information in transmission in the system. For the administration accessibility and superior, they pick virtualization innovations and apply solid confirmation and approval plots in their cloud areas. In any case, as another foundation/stage prompting new application/administration models of the future's IT industry, the necessity for a security distributed computing is not the same as the customary security issues.

A. Concerns at Different Levels

• The cloud infrastructure providers
• The cloud service providers
• The cloud consumers

Regarding data/information security, the users at different levels have variant expectations and concerns due to the roles they play in the data's life cycle. From the perspective of cloud consumers, normally who are the data owners, the concerns are essentially raised from the loss of control when the data are in a cloud. As the dataset is stored in unknown third-party infrastructure, the owner loses not only the advantages of endpoint restrictions and management, but also the fine-grained credential quality control. The uncertainty about the privacy and the doubt about the vulnerability are also resulted from the disappearing physical and logical network boundaries [36].

B. Technical and Nontechnical Challenges

The above analysis has shown that besides technical challenges, the cloud computing platform (infrastructure and service) providers are also required to meet couple of non technical issues—for example, the lack of legal requirements on data security to service providers [36]. More specifically, the following technical challenges need to be addressed in order to make cloud computing acceptable for common consumers With respect to above specialized issues, really they have been and will be tended to by steady improvement of new advances. Notwithstanding, a few extraordinary endeavors are expected to address the nontechnical difficulties. For example, a standout amongst the most troublesome issue to be illuminated in distributed computing is the clients' dread of losing authority over their information. Since end clients feel that they don't obviously know where and how their information are dealt with, or when the clients understand that their information are prepared, transmitted, and put away by gadgets under the control of a few outsiders, it is sensible for them to be worried about things occurring in the cloud. In customary workplaces, so as to keep a dataset secure, the administrator just wards off it from the danger. In cloud figuring, in any case, it appears that datasets are drawn nearer to their dangers; that is, they are transmitted to, put away in, and controlled by remote gadgets constrained by outsiders, not by the proprietor of the informational index. It is perceived this is halfway a mental issue; yet until end clients have enough data and understanding that influence them to trust distributed computing security and its elements, the dread is probably not going to leave

VI. CONCLUSION

In this section we have introduced the best in class inquire about advancement and results of secure appropriated information stockpiling in distributed computing. Distributed computing has obtained extensive consideration from both industry and the scholarly community as of late. Among all the major building squares of distributed computing, information stockpiling plays a very imperative job. Right now, there are a few difficulties in actualizing appropriated capacity in distributed computing conditions. These difficulties should be tended to before clients can appreciate the full focal points of distributed computing. In expansion, security is dependably a huge issue in any registering framework. Therefore, we studied a number of points related to the testing issues of verifying dispersed information stockpiling, including database redistributing and inquiry uprightness confirmation, information honesty in dishonest capacity, Web-application-based security, also, sight and sound information security.

REFERENCES

[1] J. A. Garay, R. Gennaro, C. Jutla, and T. Rabin, Secure distributed storage and retrieval, in Proceedings of the 11th International workshop on Distributed Algorithms, Saarbrucken, pp. 275_289 Germany, September 1997.

[2] V. Kher and Y. Kim, Securing distributed storage: Challenges, techniques, and systems, in Proceedings of the 2005 ACM Workshop on Storage Security and Survivability, Fairfax, VA, November 11, 2005.

[3] R. Ranjan, A. Harwood, and R. Buyya, Peer-to-peer-based resource discovery in global grids: A tutorial, IEEE Communications Surveys & Tutorials, 10(2), 2008, pp. 6_33.

[4] K. M. Khan, Security dynamics of cloud computing, Cutter IT Journal, June/July 2009, pp. 38_43.

[5] J. Heiser and M. Nicolett, Assessing the Security Risks of Cloud Computing, Gartner Inc., June 2, 2008.

[6] G. A. Gibson and R. V. Meter, Network attached storage architecture, Communications of the ACM, 43(11): 37_45, 2000.